METHODS IN BRIEF

GENOMICS

Variants in a bottle

Calling genetic variants from human genomic data is a challenge: many sequence differences result from systematic errors that cannot be removed simply by sequencing the same bases multiple times. Sophisticated tools exist to filter errors, but an adequate benchmark to test performance is lacking. Zook *et al.* with the US National Institute of Standards and Technology and the Genome in a Bottle Consortium redress the situation by providing a reference variant set for NA12878, an individual in the HapMap and 1000 Genomes Project. They call variants using five sequencing platforms, seven read mappers and three variant callers to remove platform and analysis-specific biases. To integrate the data sets, they enhance existing software to arbitrate calls between data sets that disagree. Software developers and geneticists can now sequence reference material from NA12878 to rigorously assess the quality of their variant-calling pipelines.

Zook, J.M. *et al. Nat. Biotechnol.* **32**, 246–251 (2014).

STEM CELLS

Efficient reprogramming from a privileged cell state

Reprogramming of somatic cells to induced pluripotency is in most cases very inefficient. It is thought to be a stochastic process, with as-yet unknown factors influencing the emergence of rare cells in which the barriers for cell fate conversion are removed or much lowered. Guo *et al.* now identify in mouse cells a 'privileged' state for which reprogramming is fast, synchronous and efficient and that therefore does not have the hallmarks of a stochastic process at all. Working first with mouse granulocyte-monocyte progenitors, which are known to reprogram at high efficiency, but then extending their studies to other murine hematopoietic cell populations as well as mouse embryonic fibroblasts, the researchers identify a very fast cell-cycle duration as a prospective indicator of the privileged, reprogrammable state.

Guo, S. *et al. Cell* **156**, 649–662 (2014).

BIOPHYSICS

Better estimates of diffusion

A ubiquitous method for estimating the diffusion rate of a particle follows directly from Einstein's observation that particle displacement increases over time in proportion to the diffusion constant and involves simple fitting of a straight line to measured values of the mean-square displacement. Vestergaard *et al.* show that estimates obtained from a single trajectory using this method are intrinsically correlated and that the precision of an estimated diffusion constant decreases as more data are used to fit the line. They present an alternative estimator that performs better for typical single-molecule data and estimated the diffusion constant of a DNA repair protein moving on DNA. The method should be extendable to two-dimensional diffusion data. Researchers measuring diffusion using a single or limited numbers of particle tracks are advised to examine this alternative estimator.

Vestergaard, C.L. *et al. Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **89**, 022726 (2014).

GENOMICS

miRNA profiling depends on platform

MicroRNAs (miRNAs) regulate diverse processes in the cell by tuning the levels of target RNAs. By comparing miRNA expression between tumor and matched nontumor tissue, scientists can identify miRNAs that play a role in cancer. Two common profiling methods are miRNA microarrays and high-throughput miRNA sequencing. A report by Wan *et al.* indicates that results from these platforms should be interpreted with caution. The researchers find very low concordance between array and sequencing data on identical ovarian cancer samples from The Cancer Genome Atlas. Moreover, miRNAs associated with survival in ovarian cancer do not overlap between the two platforms. The causes of the discordance have yet to be identified.

Wan, Y.-W. *et al. PLoS One* **9**, e87782 (2014).